RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/089,292B_

Source: _1 Fw/6_

Date Processed by STIC: _1/3/05_

ENTERED

IFW16

```
 3 <110> APPLICANT: The University of Hong Kong
 4        Ng, Hon Mun
 6 <120> TITLE OF INVENTION: Novel HEV Antigenic Peptide and Methods
 8 <130> FILE REFERENCE: 8737-000010
10 <140> CURRENT APPLICATION NUMBER: US 10/089,292B
11 <141> CURRENT FILING DATE: 2002-08-28
13 <150> PRIOR APPLICATION NUMBER: PCT/IB00/01393
14 <151> PRIOR FILING DATE: 2000-09-28
16 <150> PRIOR APPLICATION NUMBER: CA 2,283,538
17 <151> PRIOR FILING DATE: 1999-09-30
19 <160> NUMBER OF SEQ ID NOS: 18
21 <170> SOFTWARE: PatentIn version 3.3
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 642
25 <212> TYPE: DNA
26 <213> ORGANISM: Hepatitis E virus
29 <220> FEATURE:
30 <221> NAME/KEY: CDS
31 <222> LOCATION: (1)..(642)
33 <400> SEQUENCE: 1
34 cag ctg ttc tac tct cgt ccc gtc gtc tca gcc aat ggc gag ccg act      48
35 Gln Leu Phe Tyr Ser Arg Pro Val Val Ser Ala Asn Gly Glu Pro Thr
36 1               5                   10                  15
38 gtt aag ctt tat aca tct gta gag aat gct cag cag gat aag ggt att      96
39 Val Lys Leu Tyr Thr Ser Val Glu Asn Ala Gln Gln Asp Lys Gly Ile
40             20                  25                  30
42 gca atc ccg cat gac atc gac ctc ggg gag tct cgt gta gtt att cag     144
43 Ala Ile Pro His Asp Ile Asp Leu Gly Glu Ser Arg Val Val Ile Gln
44         35                  40                  45
46 gat tat gac aac caa cat gag cag gac cga ccg aca cct tcc cca gcc     192
47 Asp Tyr Asp Asn Gln His Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala
48     50                  55                  60
50 cca tcg cgc cct ttt tct gtc ctc cga gct aat gat gtg ctt tgg ctt     240
51 Pro Ser Arg Pro Phe Ser Val Leu Arg Ala Asn Asp Val Leu Trp Leu
52 65                  70                  75                  80
54 tct ctc acc gct gcc gag tat gac cag tcc act tac ggc tct tcg acc     288
55 Ser Leu Thr Ala Ala Glu Tyr Asp Gln Ser Thr Tyr Gly Ser Ser Thr
56             85                  90                  95
58 ggc cca gtc tat gtc tct gac tct gtg acc ttg gtt aat gtt gcg acc     336
59 Gly Pro Val Tyr Val Ser Asp Ser Val Thr Leu Val Asn Val Ala Thr
60             100                 105                 110
62 ggc gcg cag gcc gtt gcc cgg tca ctc gac tgg acc aag gtc aca ctt     384
63 Gly Ala Gln Ala Val Ala Arg Ser Leu Asp Trp Thr Lys Val Thr Leu
```

RAW SEQUENCE LISTING   DATE: 01/03/2005
PATENT APPLICATION: US/10/089,292B   TIME: 15:49:58

Input Set : D:\8737_000010.ST25.txt
Output Set: N:\CRF4\01032005\J089292B.raw

```
64          115                 120                 125
66 gat ggt cgc ccc ctt tcc acc atc cag cag tat tca aag acc ttc ttt      432
67 Asp Gly Arg Pro Leu Ser Thr Ile Gln Gln Tyr Ser Lys Thr Phe Phe
68       130                 135                 140
70 gtc ctg ccg ctc cgc ggt aag ctc tcc ttt tgg gag gca ggt act act      480
71 Val Leu Pro Leu Arg Gly Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr
72 145                 150                 155                 160
74 aaa gcc ggg tac cct tat aat tat aac acc act gct agt gac caa ctg      528
75 Lys Ala Gly Tyr Pro Tyr Asn Tyr Asn Thr Thr Ala Ser Asp Gln Leu
76             165                 170                 175
78 ctc gtt gag aat gcc gct ggg cat cgg gtt gct att tcc act tac acc      576
79 Leu Val Glu Asn Ala Ala Gly His Arg Val Ala Ile Ser Thr Tyr Thr
80             180                 185                 190
82 act agc ctg ggt gct ggt ccc gtc tct att tcc gcg gtt gct gtt tta      624
83 Thr Ser Leu Gly Ala Gly Pro Val Ser Ile Ser Ala Val Ala Val Leu
84             195                 200                 205
86 gcc ccc cct ccg cgc tag                                              642
87 Ala Pro Pro Pro Arg
88       210
91 <210> SEQ ID NO: 2
92 <211> LENGTH: 213
93 <212> TYPE: PRT
94 <213> ORGANISM: Hepatitis E virus
96 <400> SEQUENCE: 2
98 Gln Leu Phe Tyr Ser Arg Pro Val Val Ser Ala Asn Gly Glu Pro Thr
99 1               5                   10                  15
102 Val Lys Leu Tyr Thr Ser Val Glu Asn Ala Gln Gln Asp Lys Gly Ile
103             20                  25                  30
106 Ala Ile Pro His Asp Ile Asp Leu Gly Glu Ser Arg Val Val Ile Gln
107             35                  40                  45
110 Asp Tyr Asp Asn Gln His Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala
111      50                  55                  60
114 Pro Ser Arg Pro Phe Ser Val Leu Arg Ala Asn Asp Val Leu Trp Leu
115 65                  70                  75                  80
118 Ser Leu Thr Ala Ala Glu Tyr Asp Gln Ser Thr Tyr Gly Ser Ser Thr
119             85                  90                  95
122 Gly Pro Val Tyr Val Ser Asp Ser Val Thr Leu Val Asn Val Ala Thr
123             100                 105                 110
126 Gly Ala Gln Ala Val Ala Arg Ser Leu Asp Trp Thr Lys Val Thr Leu
127      115                 120                 125
130 Asp Gly Arg Pro Leu Ser Thr Ile Gln Gln Tyr Ser Lys Thr Phe Phe
131      130                 135                 140
134 Val Leu Pro Leu Arg Gly Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr
135 145                 150                 155                 160
138 Lys Ala Gly Tyr Pro Tyr Asn Tyr Asn Thr Thr Ala Ser Asp Gln Leu
139             165                 170                 175
142 Leu Val Glu Asn Ala Ala Gly His Arg Val Ala Ile Ser Thr Tyr Thr
143             180                 185                 190
146 Thr Ser Leu Gly Ala Gly Pro Val Ser Ile Ser Ala Val Ala Val Leu
```

```
147         195             200             205
150 Ala Pro Pro Pro Arg
151      210
154 <210> SEQ ID NO: 3
155 <211> LENGTH: 34
156 <212> TYPE: DNA
157 <213> ORGANISM: Artificial
159 <220> FEATURE:
160 <223> OTHER INFORMATION: Cloning Primer ORF2Rb
162 <400> SEQUENCE: 3
163 ggcgaatccc tagcgcggag gggggctaa aaca                              34
166 <210> SEQ ID NO: 4
167 <211> LENGTH: 2054
168 <212> TYPE: DNA
169 <213> ORGANISM: Hepatitis E virus
171 <400> SEQUENCE: 4
172 atgcgccctc ggcctatttt gctgttgctc ctcatgtttc tgcctatgct gcccgcgcca    60
174 ccgcccggtc agccgtctgg ccgccgtcgt gggcggcgca gcggcggttc cggcggtggt   120
176 ttctgggggtg accgggttga ttctcagccc ttcgcaatcc cctatattca tccaaccaac   180
178 cccttcgccc ccgatgtcac cgctgcggcc ggggctggac ctcgtgttcg ccaacccgcc   240
180 cgaccactcg gctccgcttg gcgtgaccag gcccagcgcc cgccgttgc ctcacgtcgt    300
182 agacctacca cagctggggc cgcgccgcta accgcggtcg ctccggccca tgacaccccg   360
184 ccagtgcctg atgttgactc ccgcggcgcc atcctgcgcc ggcagtataa cctatcaaca   420
186 tctcccctta cttcttccgt ggccaccggt acaaacttgg ttctatacgc cgctcctctt   480
188 agcccacttc tacccctcca ggacggcacc aatactcata taatggccac agaagcttct   540
190 aattatgccc agtaccgggt tgctcgtgcc acaattcgct accgcccgct ggtccccaac   600
192 gctgttggtg gctacgccat ctccatctcg ttctggccac agaccaccac caccccgacg   660
194 tccgttgaca tgaattcaat aacctcgacg gatgttcgta ttttagtcca gcccggcata   720
196 gcctccgagc ttgttatccc aagtgagcgc gtacactacc gtaaccaagg ttggcgctct   780
198 gttgagacct ccggggtggc ggaggaggag gccacctctg tcttgttat gctctgcata    840
200 catggctcac ctgtaaattc ttatactaat acaccttata ccggtgccct cgggctgttg   900
202 gactttgccc tcgaacttga gttccgcaac ctcacccccg gtaataccaa cacgcgggtc   960
204 tcccgttact ccagcactgc ccgtcaccgc cttcgtcgcg gtgcagatgg gactgccgag  1020
206 cttaccacca cggctgctac ccgcttcatg aaggacctct attttactag tactaatggt  1080
208 gtcggtgaga tcggccgtgg gatagcgctt accctgttta accttgctga caccctgctt  1140
210 ggcggtctac cgacagaatt gatttcgtcg gctggtggcc agctgttcta ctctcgtccc  1200
212 gtcgtctcag ccaatggcga gccgactgtt aagctttata catctgtaga gaatgctcag  1260
214 caggataagg gtattgcaat cccgcatgac atcgacctcg gggagtctcg tgtagttatt  1320
216 caggattatg acaaccaaca tgagcaggac cgaccgacac cttccccagc cccatcgcgc  1380
218 cctttttctg tcctccgagc taatgatgtg ctttggcttt ctctcaccgc tgccgagtat  1440
220 gaccagtcca cttacggctc ttcgaccggc ccagtctatg tctctgactc tgtgaccttg  1500
222 gttaatgttg cgaccggcgc gcaggccgtt gcccggtcac tcgactggac caaggtcaca  1560
224 cttgatggtc gcccccttc caccatcaag cagtattcaa agaccttctt tgtcctgccg  1620
226 ctccgcggta agctctcctt ttgggaggca ggtactacta aagccgggta cccttataat  1680
228 tataacacca ctgctagtga ccaactgctc gttgagaatg ccgctgggca tcgggttgct  1740
230 atttccactt acaccactag cctgggtgct ggtcccgtct ctatttccgc ggttgctgtt  1800
232 ttagcccccc actccgcgct agcattgctt gaggatacca tggactaccc tgcccgcgcc  1860
234 catactttcg atgacttctg tcccggagtgc cgccccttg gcctccaggg ctgtgctttt  1920
236 cagtctactg tcgctgagct tcagcgcctt aagatgaagg tgggtaaaac tcgggagtta  1980
```

RAW SEQUENCE LISTING   DATE: 01/03/2005
PATENT APPLICATION: US/10/089,292B   TIME: 15:49:58

Input Set : D:\8737_000010.ST25.txt
Output Set: N:\CRF4\01032005\J089292B.raw

```
238 tagtttattt gcttgtgccc cccttctttc tgttgcttat ttctcttttc tgcgttccgc    2040
240 gctccctgaa aaaa                                                      2054
243 <210> SEQ ID NO: 5
244 <211> LENGTH: 370
245 <212> TYPE: DNA
246 <213> ORGANISM: Hepatitis E virus
248 <400> SEQUENCE: 5
249 tgaataacat gtcttttgct gcgcccatgg gttcgcgacc atgcgccctc ggcctatttt      60
251 gctgttgctc ctcatgtttc tgcctatgct gcccgcgcca ccgcccggtc agccgtctgg     120
253 ccgccgtcgt gggcggcgca gcggcggttc cggcggtggt ttctggggtg accgggttga     180
255 ttctcagccc ttcgcaatcc cctatattca tccaaccaac cccttcgccc cgatgtcacc     240
257 gctgcggccg gggctggacc tcgtgttcgc caacccgccc gaccactcgg ctccgcttgg     300
259 cgtgaccagg cccagcgccc cgccgttgcc tcacgtcgta gacctaccac agctggggcc     360
261 gcgccgctaa                                                            370
264 <210> SEQ ID NO: 6
265 <211> LENGTH: 114
266 <212> TYPE: DNA
267 <213> ORGANISM: Hepatitis E virus
270 <220> FEATURE:
271 <221> NAME/KEY: CDS
272 <222> LOCATION: (1)..(114)
274 <400> SEQUENCE: 6
275 gac ctc gtg ttc gcc aac ccg ccc gac cac tcg gct ccg ctt ggc gtg       48
276 Asp Leu Val Phe Ala Asn Pro Pro Asp His Ser Ala Pro Leu Gly Val
277 1               5                   10                  15
279 acc agg ccc agc gcc ccg ccg ttg cct cac gtc gta gac cta cca cag       96
280 Thr Arg Pro Ser Ala Pro Pro Leu Pro His Val Val Asp Leu Pro Gln
281             20                  25                  30
283 ctg ggg ccg cgc cgc taa                                              114
284 Leu Gly Pro Arg Arg
285         35
288 <210> SEQ ID NO: 7
289 <211> LENGTH: 37
290 <212> TYPE: PRT
291 <213> ORGANISM: Hepatitis E virus
293 <400> SEQUENCE: 7
295 Asp Leu Val Phe Ala Asn Pro Pro Asp His Ser Ala Pro Leu Gly Val
296 1               5                   10                  15
299 Thr Arg Pro Ser Ala Pro Pro Leu Pro His Val Val Asp Leu Pro Gln
300             20                  25                  30
303 Leu Gly Pro Arg Arg
304         35
307 <210> SEQ ID NO: 8
308 <211> LENGTH: 22
309 <212> TYPE: DNA
310 <213> ORGANISM: Artificial
312 <220> FEATURE:
313 <223> OTHER INFORMATION: RT Primer E3R
315 <400> SEQUENCE: 8
```

RAW SEQUENCE LISTING   DATE: 01/03/2005
PATENT APPLICATION: US/10/089,292B   TIME: 15:49:58

Input Set : D:\8737_000010.ST25.txt
Output Set: N:\CRF4\01032005\J089292B.raw

```
316 cggggagtca acatcaggca ct                                      22
319 <210> SEQ ID NO: 9
320 <211> LENGTH: 24
321 <212> TYPE: DNA
322 <213> ORGANISM: Artificial
324 <220> FEATURE:
325 <223> OTHER INFORMATION: RT Primer E5R
327 <400> SEQUENCE: 9
328 aagcaaataa actataactc ccga                                    24
331 <210> SEQ ID NO: 10
332 <211> LENGTH: 34
333 <212> TYPE: DNA
334 <213> ORGANISM: Artificial
336 <220> FEATURE:
337 <223> OTHER INFORMATION: Cloning Primer ORF2F
339 <400> SEQUENCE: 10
340 gctggatccc agctgttcta ctctcgtccc gtcg                         34
343 <210> SEQ ID NO: 11
344 <211> LENGTH: 30
345 <212> TYPE: DNA
346 <213> ORGANISM: Artificial
348 <220> FEATURE:
349 <223> OTHER INFORMATION: Cloning Primer ORF2Ra
351 <400> SEQUENCE: 11
352 ggcgaattcc aaataaacta taactcccga                              30
355 <210> SEQ ID NO: 12
356 <211> LENGTH: 30
357 <212> TYPE: DNA
358 <213> ORGANISM: Artificial
360 <220> FEATURE:
361 <223> OTHER INFORMATION: Cloning Primer ORF3F
363 <400> SEQUENCE: 12
364 ccgggatccg acctcgtgtt cgccaacccg                              30
367 <210> SEQ ID NO: 13
368 <211> LENGTH: 31
369 <212> TYPE: DNA
370 <213> ORGANISM: Artificial
372 <220> FEATURE:
373 <223> OTHER INFORMATION: Cloning Primer ORF3R
375 <400> SEQUENCE: 13
376 caggaattcc ttagcggcgc ggccccagct g                            31
379 <210> SEQ ID NO: 14
380 <211> LENGTH: 21
381 <212> TYPE: DNA
382 <213> ORGANISM: Artificial
384 <220> FEATURE:
385 <223> OTHER INFORMATION: PCR Primer A3R
387 <400> SEQUENCE: 14
388 ggctcaccgg agtgtttctt c                                       21
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 01/03/2005
PATENT APPLICATION: US/10/089,292B   TIME: 15:49:59

Input Set : D:\8737_000010.ST25.txt
Output Set: N:\CRF4\01032005\J089292B.raw

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:3,8,9,10,11,12,13,14,15,16,17

VERIFICATION SUMMARY   DATE: 01/03/2005
PATENT APPLICATION: **US/10/089,292B**   TIME: 15:49:59

Input Set : **D:\8737_000010.ST25.txt**
Output Set: **N:\CRF4\01032005\J089292B.raw**